CLAIMS

Having described and disclosed my invention, I claim:

1. A document management system that comprises:

One. or a plurality of parsing engines that parse documents contained in data streams fed into said parsing engines. Each one of the parsing engines may contain mapping programs to map each page of an input document into one or a plurality of predetermined presentation zones following the presentation template, parsing programs to parse the presentation of each presentation zone into one or a plurality of output presentation formats determined by the presentation template, and reconstructing programs to reconstruct the completed presentation of said page in one or a plurality presentation options by retrieving and placing selected output presentation formats called for by the specific presentation option or options into each presentation zone following the presentation template.

One or a plurality of repository engines to store the output data sent from the parsing engines.

2. A document management system of claim 1 wherein a parsing engine may retrieve the output data of a document parsed by another parsing engine and reconstruct said document.

3. A document management system of claim 2 wherein the parsing process of a

page comprises one or a plurality of parsing layers. In each parsing layer, the mapping program maps the page into parsing zones to be parsed, following the sub-layer template designed for said sub-layer, and all parsing zones are parsed with the parsing program called for by the sub-layer template.

4. A document management system of claim 3 wherein the reconstruction process of the complete presentation of a page may include the superimposing of all reconstructed sub-layer presentations.

5. A document management system of claim 2 wherein the output presentation format options comprise:

- Figure in image presentation format.
- Text in text format describing the text content. In this format, the recognition of text symbols is required, and the text can be regenerated with different fonts and sizes.
- Mathematic equation presentation in executable equation format.
- Spreadsheet presentation in executable calculation spreadsheet format.
- Blank presentation format showing blank
- File inserting format for inserting pre-generated file.

6. A document management system of claim 5 wherein the presentation zone categories may include:

- Static figure in image presentation format zone wherein the figure images

21

are unchanged with every document of the same type and are presented in image presentation formats.

- Dynamic figure in image presentation format zone wherein the figure images may vary with every document and are presented in image presentation formats.

- Static text in text format presentation zone wherein the original text contents remain unchanged for every document of the same type and must be preserved during the parsing process. The text is presented in text format and can be regenerated with different text fonts and sizes.

- Dynamic text in text format presentation zone wherein the original text contents may vary with every document and must be preserved during the parsing process. The text is presented in text format and can be regenerated with different text fonts and sizes.

- Executable equation format zone wherein text and standard mathematical symbols such as +, -, x, /, =, and % form executable equations.

- Executable calculation spreadsheet format zone wherein data is formed into columns and rows of spreadsheets that can perform executable calculations.

- No parsing with blank presentation zone wherein the output presentation is blanked.

- No parsing with file inserting format zone wherein the output presentation is reconstructed with pre-generated data files.

7. A document management system of claim 6 wherein the parsing process further comprises document resume parsing process, which parses the resume zones containing the document key descriptive attributes into text format presentation following the resume template. The output data is stored as resume file.

8. A document management system of claim 7 wherein the parsing process further comprises document indexing parsing process, which parses the indexing attribute zones containing the document key indexing attributes into text format presentation following the indexing template. The output data is stored as indexing file.

9. A document management system of claim 8 wherein the document indexing attributes may include the organization identification, the document identification, the user identification, the account identification, and the document type identification.

10. A document management system of claim 9 wherein the parsing engines may contain the presentation templates, the resume templates and the indexing template of documents to be parsed by said parsing engine. The indexing template is identical for all documents. The presentation templates and the resume templates vary with the organizations and the document types, and may be stored with two indexing attributes: the organization identification and the

document type identification.

11. A document management system of claim 10 wherein the parsing process comprises:

- Load the indexing template.

- Parse the document indexing attribute zones to identify the organization identification and the document type identification.

- Load the resume template.

- Parse the resume zones, collect, and store the resume data.

- Load the presentation template.

- Parse the presentation zones, collect, and store the presentation data.

12. A document management system of claim 11 wherein the image reconstruction process comprises:

- Load the presentation template.

- Retrieve the presentation data and reconstruct the document image following the presentation template.

13. The document management system of claim 8 wherein the repository engine may contain three directories: the indexing directory that stores the indexing files, the resume directory that stores the resume files and the document directory that stores the document files containing the output document presentation data.

14. The document management system of claim 13 wherein the storage

hierarchy of the indexing directory is the organization identification, the user identification, and the file name is the document identification.

15. The document management system of claim 13 wherein the storage hierarchy of the resume and document directories is the organization identification, and the file name is the document identification.

16. The document management system of claim 14 wherein the storage hierarchy of the resume and document directories is the organization identification, and the file name is the document identification.

17. The document management system of claim 16 wherein the document file searching process comprises:

- Enter the indexing directory
- Using two indexing attributes: the organization identification and the user identification, search and open the indexing data of all documents belong to the said user.
- From the indexing data of indexing files, using other indexing attributes such as document type identification, account identification and date to determine the specific file that is searched for, and identify the document identification for the next step.
- Enter the resume and document directories.
- Search and retrieve the resume and document files using two indexing

attributes: the organization identification and the document identification.

18. The document management system of claim 16 wherein the document file searching process comprises:

- The user inputs the organization identification and the user identification.

- The repository engine enters the indexing directory, retrieves the indexing attributes data of all documents belong to said user in said organization, constructs, and displays an indexing table, which contains all indexing attributes of all document in an organized fashion following a predetermined grouping hierarchy.

- The user selects the document using other indexing attributes.

- The repository engine enters the resume and document directories and uses the organization identification and document identification to locate and retrieve the resume and document files of the document selected by the user.

19. A document management system of claim 9 wherein the static figures and text in image presentation format are generated in advance and stored in the parsing engines with two indexing attributes: organization identification and document type identification and the zones containing said static figures and text are designated as No parsing with file inserting format by the presentation templates. During the reconstruction process, the static figures and text are retrieved and inserted back into said zones following the presentation templates.

26

20. A document management system of claim 6 wherein the zones containing the document key descriptive attributes are parsed into text format presentation.

21. A document management system of claim 6 wherein the zones containing the document key indexing attributes are parsed into text format presentation.

22. A document management system of claim 21 wherein the zones containing the document key descriptive attributes are parsed into text format presentation.

23. A document management system of claim 1 which further comprises one or a plurality of document acquisition engines that collect document presentation data from document generators, convert data into one single format, envelop data into data streams and send said data streams to the parsing engines.

24. A document management service provider comprising:

A host document management system that is locally located at the facility of said service provider. The said host document management system comprises one or a plurality of Web server engines that facilitate the communication services such as receiving and responding to the user requests between the host document management system and a collection of Internet user systems, security engines for authenticating users and authorizing services, host parsing engines for parsing documents, host repository engines for storing output data

27

and host communication engines for controlling the communication between the host and the client systems.

One or a plurality of client document management systems that is remotely located at the facilities of the service subscribers and is linked to the host system across the Internet. Each of the said client document management systems comprises client parsing engines for parsing documents, client repository engines for storing output data from the client parsing engines, client communication engines for controlling the communication between the host system and the said client system and document acquisition engines to collect document presentation data from a collection of client document generators and computers; the document acquisition engines convert data into one single format, envelop data into data streams and send said data streams to the client parsing engines. The parsed data is sent across the Internet to the host document management system where the documents are reconstructed by the host parsing engines. Data is stored at either the client repository engines or the host repository engines or both.

One or a plurality of client document acquisition engines those are remotely located at the facilities of the service subscribers. The document acquisition engines collect document presentation data from a collection of client document generators and computers, convert data into one single format, envelop data into data streams and send said data streams across the Internet to the host system where the documents are parsed, and reconstructed by the host parsing engines. Data is stored at the host repository engines.